

RESEARCH

Long term duration of protective effect for HPV negative women: follow-up of primary HPV screening randomised controlled trial

 OPEN ACCESS

K Miriam Elfström *postgraduate student*¹, Vitaly Smelov *research fellow*², Anna L V Johansson *statistician*¹, Carina Eklund *research coordinator*², Pontus Naucclér *research fellow*^{3,4}, Lisen Arnheim-Dahlström *associate professor*¹, Joakim Dillner *professor*^{1,2}

¹Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Box 281, 171 77 Stockholm, Sweden; ²Department of Laboratory Medicine, Tumor and Cell Biology, Karolinska Institutet, Stockholm, Sweden; ³Department of Microbiology, Tumor and Cell Biology, Karolinska Institutet, Stockholm, Sweden; ⁴Department of Infectious Diseases, Karolinska University Hospital, Stockholm, Sweden

Abstract

Objectives To assess whether the increased sensitivity of screening for human papillomavirus (HPV) may represent overdiagnosis and to compare the long term duration of protective effect against cervical intraepithelial neoplasia grade 2 or worse (CIN2+) in HPV based and cytology based screening.

Design 13 year follow-up of the Swedescreen randomised controlled trial of primary HPV screening.

Setting Organised cervical screening programme in Sweden.

Participants 12 527 women aged 32-38 attending organised screening were enrolled and randomised to HPV and cytology double testing (intervention arm, n=6257) or to cytology only, with samples frozen for future HPV testing (control arm, n=6270).

Main outcome measures Cumulative incidence of CIN2+ and CIN3+ (Kaplan Meier curves). Longitudinal test characteristics were calculated for cytology only, HPV testing only, and cytology and HPV testing combined, adjusting for censoring.

Results The increased detection of CIN2+ in the intervention arm decreased over time. After six years, the cumulative incidence of CIN3+ was similar in both trial arms, and after 11 years the cumulative incidence of CIN2+ became similar in both arms. The longitudinal sensitivity of cytology for CIN2+ in the control arm at three years was similar to the sensitivity of HPV testing in the intervention arm at five years of follow-up: 85.94% (95% confidence interval 76.85% to 91.84%) v 86.40% (79.21% to 91.37%). The sensitivity of HPV screening for CIN3+ after five years was 89.34% (80.10% to 94.58%) and for cytology after three years was 92.02% (80.59% to 96.97%).

Conclusions Over long term follow-up, the cumulative incidence of CIN2+ was the same for HPV screening and for cytology, implying that the increased sensitivity of HPV screening for CIN2+ reflects earlier detection rather than overdiagnosis. The low long term risks of CIN3+ among women who tested negative in HPV screening, support screening intervals of five years for such women.

Trial registration Clinicaltrials.gov NCT00479375.

Introduction

Cervical screening using cytology has resulted in a noticeable reduction in cervical cancer. However, there are still about 55 000 annual cases of cervical cancer in the European region.¹ Infection with oncogenic types of human papillomavirus (HPV) is a necessary risk factor in cervical carcinogenesis² and testing for HPV DNA has a higher sensitivity for detection of cervical intraepithelial neoplasia, the precursor lesion of cervical cancer.³⁻⁶ Therefore HPV based cervical screening might possibly enable screening programmes with an increased protective effect against cervical cancer. Because HPV infection precedes the development of cervical intraepithelial neoplasia,⁷ it is conceivable that HPV based screening could be performed with longer screening intervals, which in turn could result in more cost effective screening.⁸

Randomised controlled trials of HPV based screening have found that the initial higher detection of cervical intraepithelial neoplasia grade 3 or worse (CIN3+) is followed by a reduced detection rate in subsequent screening rounds.^{4,9-11} The fact that the number of lesions detected at baseline HPV screening, in

Correspondence to: J Dillner joakim.dillner@ki.se

Extra material supplied by the author (see <http://www.bmj.com/content/348/bmj.g130?tab=related#webextra>)

Figure showing cumulative incidence of CIN2+
Figure showing cumulative incidence of CIN3+

particular for CIN2, is greater than in subsequent screens^{4 10 12} has raised concerns that the increased detection rate may represent overdiagnosis of lesions that would not have progressed to invasive cancer. However, it is also possible that the increased detection rate may represent early detection. Follow-up studies encompassing more screening rounds could assist in distinguishing between these possibilities.

The risk of CIN3+ among women screening test negative was found to be about the same after six years in HPV screening as after three years in cytology screening,¹³ suggesting that screening intervals could safely be extended for women who screen test negative for HPV. Cohort studies have found that HPV negative women continue to have low risks for 10 years,^{14 15} but randomised controlled trials have so far only reported results up to six years.⁹ The Swedescreen randomised HPV screening trial was started in 1997 and explored the effect of a single lifetime HPV test as an add-on to an organised population based cervical screening programme.⁴ Because this randomised controlled trial now has a 13 year follow-up time and since only a single HPV test was done, it is possible both to ascertain the long term duration of the protective effect against CIN2+ and CIN3+ for screen test negative women and to investigate whether the increased detection rates for CIN2+ seen at the baseline HPV screening represent overdiagnosis or merely earlier diagnosis.

Methods

The Swedescreen randomised controlled trial⁴ invited women aged 32 to 38 during 1997-2000 attending organised cervical screening in five different regions of Sweden to participate, following informed consent. The screening programme invites women to screening at three year intervals between ages 23 and 50 and five year intervals between ages 51 and 60 after a “sorting out” procedure based on comprehensive cytology/pathology registries, which register all cervical smears and biopsy samples taken in Sweden (both organised and opportunistic). Women invited for screening are chosen from the population registry, which lists all women residing in Sweden. Women are sorted out from being invited to participate in the organised screening programme if they have had a recent opportunistic smear, as assessed by registry linkage with cytology registers. In total, 12 527 women were enrolled and randomised 1:1 to HPV and cytology double testing (intervention arm, n=6257) or cytology only, with samples frozen for future HPV testing (control arm, n=6270). Randomisation was performed using a random number generator, and randomisation codes were only released after the samples had arrived at the laboratory. HPV DNA testing in the intervention arm used GP5+/6+ polymerase chain reaction (PCR) with subsequent typing by reverse line dot blot hybridisation.¹⁶ The types included in the HPV test were: 16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, 59, 66, and 68. The method was found to correlate well to the GP5+/6+ PCR HPV testing used in large clinical validation studies in the Netherlands.¹⁷ In the control arm, the samples were frozen at -80°C and retrieved for testing about 10 years later. HPV testing in the control arm was based on the same testing principle but used improvements of the test that had been introduced in the meantime—namely, a switch from AmpliTaq PCR enzyme to AmpliTaq Gold PCR enzyme,¹⁸ a modification of the general primers to increase sensitivity,¹⁹ and typing using Luminex.¹⁹ The analysis has been validated as proficient using the World Health Organization HPV DNA genotyping proficiency panel.^{20 21}

Clinical management of the cervical smear results followed the established routines of the organised programme. For this study we used 13 years of follow-up, representing four screening rounds plus one year to account for delays in attendance and any diagnostic follow-up, to assess longitudinal risks for development of CIN2+ and CIN3+. In the intervention arm, women who were HPV positive and did not have an abnormal smear test result at study enrolment were invited for a second HPV test and cytology at least 12 months later (average attendance was actually 19 months later). Women who were persistently HPV positive were invited to colposcopy, as described.^{4 22} A similar number of random repeat cytologies and colposcopies were performed in the control arm, to avoid ascertainment bias.²² The trial was designed to be double blinded. However, blinding was discontinued for safety reasons in August 2003, when it became apparent that the risks of having an undiagnosed high grade lesion were greater than expected in HPV positive women.⁴ At the time of unblinding, three years after enrolment had finished, the first round of study colposcopies had been completed.

We followed up the study participants using registry based information from 1997. The National Quality Registry for Cervical Cancer Prevention contains 100% of information on all cervical smears and biopsy samples taken in Sweden (from organised screening as well as all other tests). Given delays in the import of files from some laboratories, the present follow-up extended to September 2011 for Gothenburg, to January 2012 for Stockholm, and to December 2011 for all other laboratories in the country. Because follow-up used information from the national registry, we were able to capture the test results for women who moved within Sweden. Women lost to follow-up in the screening registers possibly resulted from death, emigration, or other reasons for long term non-attendance.

Follow-up started at the first study test result taken at recruitment (between 1997 and 2000). End of follow-up occurred at the first histologically confirmed diagnosis of a CIN2+ or CIN3+ lesion, the last registered sampling date, or at 13 years of follow-up, whichever came first. If a woman had a CIN2 lesion and later developed a CIN3 lesion, we set censoring to the date of diagnosis of the CIN2 lesion for the analysis of CIN2+ and to the date of diagnosis of the CIN3 lesion for the analysis of CIN3+. We excluded from the analysis women with missing baseline test results and those with no follow-up test results beyond the first study test result (n=436), leaving 12 091 women (96.5%) with follow-up in the analysis. We defined baseline cytology status as the first test result in the study and categorised this as abnormal if the diagnosis was atypical squamous cells of undetermined significance or worse. We categorised baseline HPV status as HPV positive and HPV negative. The impact of acting on the HPV test was evaluated by comparing the intervention arm with the control arm.

Statistical analysis

Using 1 minus the Kaplan-Meier curves, we estimated the cumulative incidences of CIN2+ and CIN3+, with 95% pointwise confidence bands based on the Greenwood formula for the standard error. We calculated cumulative incidences stratified by study arm and by baseline test results. Longitudinal test characteristics (sensitivity, specificity, and negative and positive predictive values) were calculated with adjustment for censoring. For sensitivity and specificity we adjusted the numerator and denominator for censoring by up-weighting the numbers according to the censoring distribution over follow-up time and taking into account that censoring may depend on the test result at baseline (conditional weighting).²³ If censoring is

not taken into account, the at risk population at the time points of interest may not be representative of the censored population.²³ The negative and positive predictive values are estimated from the cumulative incidence at given time points, and hence also adjusted for censoring. Longitudinal sensitivity is defined as the sensitivity for current and future disease²⁴ and is one of the indices in the framework for evaluating potential screening strategies.²⁴ We calculated test characteristics at 3, 5, 8, and 10 years of follow-up to estimate the duration of protective effect at intervals currently recommended for cytology (3 and 5 years) and some of the lengthened intervals for HPV based screening (5, 8, and 10 years). To compare the sensitivity of cytology in the control arm at three years with the sensitivity of HPV testing in the intervention arm at 3, 5, 8, and 10 years, we used a two sample test of proportions assuming the binomial distribution. We did not adjust P values for multiple testing.

Data were compiled in SAS 9.2 and analyses performed using Stata 11.

Results

From 12 527 enrolled women, 12 091 (96.5%) had baseline cytology and at least one follow-up test (fig 1). Among the 12 091 women included in this analysis, 387 developed histologically confirmed CIN2+ and 230 women developed CIN3+ during 13 years of follow-up. Overall, there were 198 cases of CIN2+ in the intervention arm and 189 in the control arm and 119 cases of CIN3+ in the intervention arm and 111 in the control arm. The median follow-up time for the whole cohort was 10.95 years (range 0.04 to 13.00 years) for the CIN2+ outcome and 10.98 years (range 0.04 to 13.00 years) for the CIN3+ outcome.

The cumulative incidence of CIN2+ among women who were cytology negative at baseline increased steadily during the follow-up period, reaching 2.73% (95% confidence interval 2.17% to 3.44%) at 13 years in the control arm (fig 2). By comparison, the cumulative incidence of CIN2+ increased slowly for HPV negative women, reaching 1.74% (1.24% to 2.45%) at 13 years in the intervention arm. At 13 years the rates of CIN2+ in the intervention arm were similar between women who were HPV and cytology double negative and HPV negative (1.63%, 1.11% to 2.32%).

During the first 11 years of follow-up, women in the intervention arm who were cytology negative at baseline had a somewhat higher cumulative incidence of CIN2+ than cytology negative women in the control arm, reflecting the effect of the intervention on HPV test results in the intervention arm (fig 2). However, after 11 years the cumulative incidence of CIN2+ among cytology negative women no longer differed (fig 2), suggesting that the cases of CIN2+ among cytology negative women in the intervention arm reflect earlier diagnosis and not overdiagnosis. The cumulative incidence of CIN2+ among HPV negative women and women double negative for HPV and cytology were similar throughout follow-up in the control arm (fig 2; see supplementary material for graphs with confidence intervals). In the intervention arm, the cumulative incidence of CIN2+ for HPV negative and double negative women were similar at seven years of follow-up.

At 13 years of follow-up the cumulative incidence of CIN3+ among women who were cytology negative at baseline was 1.54% (1.10% to 2.15%) in the control arm and 0.89% (0.53% to 1.51%) among HPV negative women in the HPV testing arm (fig 3). Women with a double negative test result had similar rates of CIN3+ as HPV negative women, reaching a cumulative incidence of 0.84% (0.48% to 1.47%) in the intervention arm

at 13 years (fig 3). During the first six years of follow-up, the cumulative incidence of CIN3+ was greater in the intervention arm (fig 3), reflecting that women persistently positive for HPV and with negative cytology had been referred to colposcopy, resulting in additional cases of CIN3+ detected. However, after six years of follow-up the CIN3+ rates did not differ, suggesting that the additional CIN3+ cases detected are more likely to reflect early diagnosis rather than overdiagnosis.

Using CIN2+ as the outcome, the longitudinal sensitivity of cytology at three years in the control arm was similar to the sensitivity of HPV testing in the intervention arm at five years of follow-up. The sensitivity of cytology at three years in the control arm was 85.94% (95% confidence interval 76.85% to 91.84%) and the sensitivity of HPV in the intervention arm at five years was 86.40% (79.21% to 91.37%), there was no significant difference between the proportions ($P=0.8970$, table 1). The longitudinal sensitivity of cytology for CIN3+ in the control arm at three years was 92.02% (80.59% to 96.97%) and 89.34% (80.10% to 94.58%) for HPV testing in the intervention arm at five years ($P=0.4871$ for difference of proportions, table 2). The specificity of HPV testing in the intervention arm was lower than for cytology in the control arm at all time points and for both CIN2+ and CIN3+ outcomes. For CIN2+, the specificity for cytology in the control arm ranged from 98.45% to 98.50% at three and 10 years, respectively, and for HPV testing from 94.05% to 94.82% (table 1). The specificity was only slightly lower for CIN3+ compared with CIN2+ (table 2). The negative predictive values remained higher for HPV based compared with cytology based screening throughout follow-up for both CIN2+ and CIN3+. The positive predictive value was highest for cytology but increased over time for HPV testing (tables 3 and 4).

Discussion

We found that a negative test result in HPV screening is associated with a low risk for cervical intraepithelial neoplasia grade 2 or worse (CIN2+) or grade 3 or worse (CIN3+) and that this low risk is of prolonged duration. The risk was about the same at five years for women who tested negative for HPV at screening as at three years after a negative cytology screening test result. Furthermore, we found that being negative for both HPV and cytology had essentially the same protection even after 13 years of follow-up. Finally, we found that clinical follow-up of women with persistent HPV infections did not result in any increased cumulative detection of CIN2 or CIN3+ over long term follow-up. This suggests that the additional lesions detected by HPV screening do not represent overdiagnosis but rather earlier diagnosis of lesions.

The Swedish randomised controlled trial on primary HPV screening was started in 1997 and compared with previous studies reporting on the longitudinal performance of HPV screening; this study contributes a considerably longer follow-up time. Although observational cohort studies are informative about the clinical course of HPV, they do not provide information on the effect of an HPV based screening programme. The trial employed a single lifetime HPV test in a setting where essentially no other HPV testing was performed. This provided a unique opportunity for delineation of the effect of a single HPV test and a chance to study whether the increased detection of CIN2+ seen with HPV screening represents overdiagnosis or merely early diagnosis of lesions that would have been detected later in a cytology based screening programme. Because the trial was nested in a population based organised screening programme, the results should be

generalisable to real life screening situations. The infrastructure for follow-up, where all cytological and histopathological laboratories in the country export the data on all specimens analysed (both organised and non-organised) ensured an as complete follow-up as possible, with 96.5% of the participants having been followed-up with at least one additional screening test.

Limitations and other considerations

The HPV DNA test used in the intervention arm was a polymerase chain reaction test that was one of the best and most well characterised tests in 1997 (GP5+/6+ PCR). The testing in the control arm was done on frozen specimens, using an HPV test based on the same principle, but that was modified to correspond to the advances made in testing by 2012. Thus, 7.3% of the participants followed in the intervention arm were HPV positive at baseline, but 9.7% were positive at baseline in the control arm. A previous study with less than five years of follow-up found that the increased sensitivity of the modern HPV test had reduced the specificity without improving the longitudinal sensitivity for CIN2+.¹⁸ The present study also found no difference in longitudinal sensitivity for the first seven years but a tendency for improved protection with the more modern test after more than seven years of follow-up. The limited differences imply that even over 15 years improvements in assays have not substantially changed the outcome and suggest that the results would be applicable to HPV testing as practised today. Our results also suggest that longitudinal follow-up is essential when determining the optimal analytical sensitivity of HPV tests considered for use in screening programmes.

The study used routine histopathological diagnoses and therefore some misclassification of endpoints is possible, particularly for CIN2. However, compared with rereviewed diagnoses, risk estimates derived from routine histopathological diagnoses performed in the screening programme reflect the real life setting and can therefore be more directly generalisable to actual screening programmes.

The study only invited women aged 32-38 years, the age when HPV testing is proposed to have the greatest effectiveness. Given the narrow age range, our study was not able to examine possible age related differences in test characteristics.

Comparisons with other studies

Compared with the 2008 joint European cohort study, which had data for only six years of follow-up,¹³ our study focused on comparing the effect of HPV based and cytology based screening in a randomised controlled trial, and follow-up was longer. Our results from the long term follow-up of the Swedescreen randomised controlled trial are similar to the findings of the European cohort study. A Dutch study in which women who screened test negative were followed for CIN3+ outcomes and which calculated sensitivity and specificity for women who were cytology negative, HPV negative, and HPV and cytology negative²⁵ found that after the first screening round (five years), the sensitivity both for HPV alone and for HPV and cytology double testing was 92.9%, which is comparable to our study arm specific results for women who were HPV negative and HPV and cytology double negative.²⁵

Several randomised controlled trials evaluating HPV and cytology screening have shown overall improved screening performance with HPV testing, although the magnitude of this benefit, duration of follow-up, and comparisons have differed across settings. In the ARTISTIC trial (A Randomised Trial of

HPV Testing in Primary Cervical Screening), HPV testing and liquid based cytology were compared with liquid based cytology alone and the results from the first two screening rounds showed a reduction in detection of high grade lesions in the second round compared with the first in the double testing arm.²⁶ Extended follow-up of the ARTISTIC trial has since provided evidence that the negative protective effect of being HPV negative was significant over three rounds of screening.⁹ The New Technologies for Cervical Cancer screening (NTCC) and the Population Based Screening Study Amsterdam (POBASCAM) trials showed similar increases in detection of high grade lesions in the first round of screening with HPV testing followed by decreases in subsequent rounds.¹¹

The impact of one screening round with HPV testing has been shown by work done in Canada,⁶ Finland,¹² and India.²⁷ The Canadian Cervical Cancer Screening Trial (CCCaST) examined baseline sensitivity and specificity of HPV testing compared with cytology, determining that HPV testing had greater sensitivity for the detection of high grade lesions.⁶ In the first round of screening, the Finnish trial showed greater detection of cervical intraepithelial neoplasia, especially low grade lesions; however, since subsequent screening rounds have not been examined yet, further comparisons could not be made.¹² The results of a community randomised trial in India showed a significant reduction in cervical cancer incidence and mortality in the HPV testing arm compared with the group receiving standard care, showing the effect of HPV testing against invasive cancer.²⁷ Together, evidence from these trials provides an increasingly strong base in support of using HPV testing in primary screening. Our study adds evidence that the early increased detection of high grade lesions with HPV testing does not represent overdiagnosis, but rather suggests a gain in lead time. We also provide long term follow-up data suggesting that the sensitivity of screening for HPV at five years is similar to that of screening using cytology at three years.

Conclusions and policy implications

The non-significant gain in lowered risks for CIN2+ and CIN3+ of being double negative (cytology and HPV negative) implies that double testing is not likely to represent an improvement in effect, while it certainly would increase costs and lower specificity. Our data thus support the use of screening for HPV as a stand-alone primary test, with cytology restricted to triaging women who are HPV positive. The results further substantiate the potential for extending the screening interval with HPV based screening. Because the sensitivity of cytology at the currently recommended screening interval in the European Union (three years) was similar to the sensitivity of HPV testing at five years, our results suggest that a five yearly screening interval for HPV negative women is an alternative to a three yearly screening interval for cytology negative women in those aged more than 30. With several effective screening methods available, the choice of optimal tests and intervals is a balance between sensitivity, specificity, and costs, where priorities may differ between countries. Finally, our study found evidence to suggest that the increased detection of CIN2 and CIN3+ during short term follow-up of HPV based screening does not represent overdiagnosis but rather early detection. This has important implications for the evaluation of the specificity and modelling of cost effectiveness of HPV based screening programmes.

We thank Mariam Lashkariani for data linkage support, Björn Strander for help with the linkage to screening data from Gothenburg, Inger Olausson for help with the linkage to screening data from Stockholm,

What is already known on this topic

Testing for human papillomavirus (HPV) DNA detects more cases of high grade cervical intraepithelial neoplasia (CIN) than does cytology, but whether this represents overdiagnosis of spontaneously regressive lesions is unclear

The risk of CIN grade 3 or worse after six years is about the same for women testing negative for HPV DNA as after three years for women with a negative cytology test result

Data from randomised clinical trials are, however, limited

What this study adds

The increased sensitivity for high grade CIN of HPV based screening was found to reflect earlier detection rather than overdiagnosis

This follow-up study of a randomised primary HPV screening trial found that the sensitivity of HPV based screening after five years is about the same as cytology based screening after three years

These results suggest that longer screening intervals could be used in HPV based screening

and all past and present members of the Swedescreen study group including researchers, healthcare practitioners, and data managers.

Contributors: KME collected and prepared the data, conducted the statistical analysis, and drafted the paper. AJ assisted with the statistical analyses. VS and CE completed the HPV DNA testing in the control arm. PN and LAD advised on analyses, data collection, and manuscript. JD conceived the study and analyses and supervised data collection and the writing of the paper. All authors helped revise the manuscript and had access to the data. JD is the guarantor of the study and accepts full responsibility for the finished article, had access to any data, and controlled the decision to publish.

Funding: This study was supported by the PREHDICT and CoheaHr projects (EU FP7 programmes), the Swedish Cancer Society, and the Swedish Foundation for Strategic Research. The funding sources had no role in the study design; the collection, analysis, and interpretation of data; and the writing of the article and the decision to submit it for publication.

Competing interests: All authors have completed the ICMJE uniform disclosure form at www.icmje.org/coi_disclosure.pdf and claim no conflict of interest related to the submitted work. JD has received grants from Merck/SPMSD for unconditional studies and LAD has received grants from Merck/SPMSD and GlaxoSmithKline for unconditional studies. No other relationships or activities that could appear to have influenced the submitted work.

Ethical approval: The Swedescreen study was approved by the ethical review board in Stockholm, Sweden (DNR 1996/305). The long term follow-up of the Swedescreen study had additional approval from the ethical review board in Stockholm, Sweden (DNR 2012/780-32).

Data sharing: All data from the study is available from the authors.

Transparency: The lead author affirms that this manuscript is an honest, accurate, and transparent account of the study being reported; that no important aspects of the study have been omitted; and that any discrepancies from the study as planned (and, if relevant, registered) have been explained.

- Forman D, de Martel C, Lacey CJ, Soerjomataram I, Lortet-Tieulent J, Bruni L, et al. Global burden of human papillomavirus and related diseases. *Vaccine* 2012;30(Suppl 5):F12-23.
- Walboomers JM, Jacobs MV, Manos MM, Bosch FX, Kummer JA, Shah KV, et al. Human papillomavirus is a necessary cause of invasive cervical cancer worldwide. *J Pathol* 1999;189:12-9.
- Arbyn M, Ronco G, Anttila A, Meijer CJ, Poljak M, Ogilvie G, et al. Evidence regarding human papillomavirus testing in secondary prevention of cervical cancer. *Vaccine* 2012;30(Suppl 5):F88-99.
- Naucler P, Ryd W, Tornberg S, Strand A, Wadell G, Elfgrén K, et al. Human papillomavirus and Papanicolaou tests to screen for cervical cancer. *N Engl J Med* 2007;357:1589-97.
- Ronco G, Giorgi-Rossi P, Carozzi F, Confortini M, Dalla Palma P, Del Mistro A, et al. Results at recruitment from a randomized controlled trial comparing human papillomavirus testing alone with conventional cytology as the primary cervical cancer screening test. *J Natl Cancer Inst* 2008;100:492-501.
- Mayrand MH, Duarte-Franco E, Rodrigues I, Walter SD, Hanley J, Ferenczy A, et al. Human papillomavirus DNA versus Papanicolaou screening tests for cervical cancer. *N Engl J Med* 2007;357:1579-88.
- Schlecht NF, Kulaga S, Robitaille J, Ferreira S, Santos M, Miyamura RA, et al. Persistent human papillomavirus infection as a predictor of cervical intraepithelial neoplasia. *JAMA* 2001;286:3106-14.
- De Kok IM, van Rosmalen J, Dillner J, Arbyn M, Sasieni P, Iftner T, et al. Primary screening for human papillomavirus compared with cytology screening for cervical cancer in European settings: cost effectiveness analysis based on a Dutch microsimulation model. *BMJ* 2012;344:e670.
- Kitchener HC, Gilham C, Sargent A, Bailey A, Albrow R, Roberts C, et al. A comparison of HPV DNA testing and liquid based cytology over three rounds of primary cervical screening: extended follow up in the ARTISTIC trial. *Eur J Cancer* 2011;47:864-71.
- Ronco G, Giorgi-Rossi P, Carozzi F, Confortini M, Dalla Palma P, Del Mistro A, et al. Efficacy of human papillomavirus testing for the detection of invasive cervical cancers and cervical intraepithelial neoplasia: a randomised controlled trial. *Lancet Oncol* 2010;11:249-57.
- Rijkaart DC, Berkhof J, Rozendaal L, van Kemenade FJ, Bulkman NW, Heideman DA, et al. Human papillomavirus testing for the detection of high-grade cervical intraepithelial neoplasia and cancer: final results of the POBASCAM randomised controlled trial. *Lancet Oncol* 2012;13:78-88.
- Leinonen MK, Nieminen P, Lonnberg S, Malila N, Hakama M, Pokhrel A, et al. Detection rates of precancerous and cancerous cervical lesions within one screening round of primary human papillomavirus DNA testing: prospective randomised trial in Finland. *BMJ* 2012;345:e7789.
- Dillner J, Rebolj M, Birembaut P, Petry KU, Szarewski A, Munk C, et al. Long term predictive values of cytology and human papillomavirus testing in cervical cancer screening: joint European cohort study. *BMJ* 2008;337:a1754.
- Kjaer S, Hogdall E, Frederiksen K, Munk C, van den Brule A, Svare E, et al. The absolute risk of cervical abnormalities in high-risk human papillomavirus-positive, cytologically normal women over a 10-year period. *Cancer Res* 2006;66:10630-6.
- Schiffman M, Glass AG, Wentzensen N, Rush BB, Castle PE, Scott DR, et al. A long-term prospective study of type-specific human papillomavirus infection and risk of cervical neoplasia among 20,000 women in the Portland Kaiser Cohort Study. *Cancer Epidemiol Biomarkers Prev* 2011;20:1398-409.
- Forslund O, Antonsson A, Edlund K, van den Brule AJ, Hansson BG, Meijer CJ, et al. Population-based type-specific prevalence of high-risk human papillomavirus infection in middle-aged Swedish women. *J Med Virol* 2002;66:535-41.
- Jacobs MV, Snijders PJ, Voorhorst FJ, Dillner J, Forslund O, Johansson B, et al. Reliable high risk HPV DNA testing by polymerase chain reaction: an intermethod and intramethod comparison. *J Clin Pathol* 1999;52:498-503.
- Wahlstrom C, Iftner T, Dillner J, Dillner L, Swedescreen study group. Population-based study of screening test performance indices of three human papillomavirus DNA tests. *J Med Virol* 2007;79:1169-75.
- Soderlund-Strand A, Carlson J, Dillner J. Modified general primer PCR system for sensitive detection of multiple types of oncogenic human papillomavirus. *J Clin Microbiol* 2009;47:541-6.
- Eklund C, Forslund O, Wallin KL, Zhou T, Dillner J, WHO Human Papillomavirus Laboratory Network. The 2010 global proficiency study of human papillomavirus genotyping in vaccinology. *J Clin Microbiol* 2012;50:2289-98.
- Eklund C, Zhou T, Dillner J, WHO Human Papillomavirus Laboratory Network. Global proficiency study of human papillomavirus genotyping. *J Clin Microbiol* 2010;48:4147-55.
- Elfgrén K, Rylander E, Radberg T, Strander B, Strand A, Paajanen K, et al. Colposcopic and histopathologic evaluation of women participating in population-based screening for human papillomavirus deoxyribonucleic acid persistence. *Am J Obstet Gynecol* 2005;193(3 Pt 1):650-7.
- Antolini L, Valsecchi MG. Performance of binary markers for censored failure time outcome: nonparametric approach based on proportions. *Stat Med* 2012;31(11-12):1113-28.
- Adriaensens WJ, Matthei C, Buntinx FJ, Arbyn M. A framework provided an outline toward the proper evaluation of potential screening strategies. *J Clin Epidemiol* 2013;66:639-47.
- Bulkman NW, Rozendaal L, Voorhorst FJ, Snijders PJ, Meijer CJ. Long-term protective effect of high-risk human papillomavirus testing in population-based cervical screening. *Br J Cancer* 2005;92:1800-2.
- Kitchener HC, Almonte M, Thomson C, Wheeler P, Sargent A, Stoykova B, et al. HPV testing in combination with liquid-based cytology in primary cervical screening (ARTISTIC): a randomised controlled trial. *Lancet Oncol* 2009;10:672-82.
- Sankaranarayanan R, Nene BM, Shastri SS, Jayant K, Muwonge R, Budukh AM, et al. HPV screening for cervical cancer in rural India. *N Engl J Med* 2009;360:1385-94.

Accepted: 6 January 2014

Cite this as: *BMJ* 2014;348:g130

This is an Open Access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 3.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/3.0/>.

Tables

Table 1| Longitudinal test characteristics using cervical intraepithelial neoplasia grade 2 or worse (CIN2+) as outcome at 3, 5, 8, and 10 years of follow-up, by cytology, human papillomavirus (HPV) status, and cytology and HPV for study arms combined and separately

Screening test	Cumulative count of outcomes at follow-up by baseline status				Test characteristics for CIN2+			
	Intervention arm		Control arm		Sensitivity (95% CI)		Specificity (95% CI)	
	Negative at baseline	Positive at baseline	Negative at baseline	Positive at baseline	Intervention arm	Control arm	Intervention arm	Control arm
Cytology								
Baseline status*:	5844	152	5868	166				
3 years	27	64	12	73	70.31 (60.17 to 78.78)	85.94 (76.85 to 91.84)†	98.51 (98.17 to 98.79)	98.45 (98.10 to 98.73)
5 years	56	70	35	75	55.36 (46.61 to 63.79)	68.02 (58.75 to 76.05)	98.61 (98.28 to 98.88)	98.48 (98.13 to 98.76)
8 years	85	71	68	77	44.59 (37.00 to 52.44)	51.97 (43.86 to 59.99)	98.62 (98.29 to 98.89)	98.50 (98.16 to 98.79)
10 years	101	73	86	77	40.40 (33.40 to 47.82)	45.07 (37.62 to 52.74)	98.66 (98.33 to 98.93)	98.50 (98.15 to 98.78)
HPV								
Baseline status‡:	5439	427	5210	562				
3 years	7	83	9	72	92.23 (84.58 to 96.25)†	88.88 (79.99 to 94.11)	94.05 (93.42 to 94.63)	91.39 (90.65 to 92.08)
5 years	17	108	13	91	86.40 (79.21 to 91.37)†	87.49 (79.64 to 92.59)	94.47 (93.85 to 95.03)	91.70 (90.97 to 92.38)
8 years	34	121	23	113	77.30 (69.95 to 83.29)†	82.88 (75.57 to 88.35)	94.69 (94.08 to 95.24)	92.09 (91.37 to 92.75)
10 years	45	128	32	121	72.45 (65.17 to 78.71)†	78.07 (70.65 to 84.03)	94.82 (94.22 to 95.37)	92.24 (91.52 to 92.91)
Cytology and HPV								
Baseline status§:	5365	663	5132	931				
3 years	3	88	5	81	96.69 (90.25 to 98.93)	94.18 (86.76 to 97.55)	90.32 (89.54 to 91.05)	85.78 (84.87 to 86.64)
5 years	11	115	9	102	91.22 (84.84 to 95.07)	91.87 (85.11 to 95.72)	90.73 (89.97 to 91.45)	86.08 (85.18 to 86.94)
8 years	26	131	19	127	82.67 (75.79 to 87.91)	86.70 (80.09 to 91.35)	90.98 (90.22 to 91.69)	86.46 (85.56 to 87.31)
10 years	37	138	28	136	77.19 (70.16 to 82.97)	81.80 (74.91 to 87.12)	91.10 (90.34 to 91.81)	86.61 (85.71 to 87.45)

*Women with unsatisfactory baseline smears and women with no baseline smears excluded.

†P value for test of difference of proportions, sensitivity of cytology in control arm at three years versus HPV testing in the intervention arm at 3 (P=0.0481), 5 (P=0.8970), 8 (p=0.0294), and 10 years (P=0.0012).

‡Women with unknown HPV baseline test excluded.

§Women with both unknown HPV and unknown cytology baseline tests excluded.

Table 2| Longitudinal test characteristics using cervical intraepithelial neoplasia grade 3 or worse (CIN3+) as outcome at 3, 5, 8, and 10 years of follow-up, by cytology, human papillomavirus (HPV) status, and cytology and HPV for study arms combined and separately

Screening test	Cumulative count of outcomes at follow-up by baseline status				Test characteristics for CIN3+			
	Intervention arm		Control arm		Sensitivity		Specificity	
	Negative at baseline	Positive at baseline	Negative at baseline	Positive at baseline	Intervention arm	Control arm	Intervention arm	Control arm
Cytology								
Baseline status*:	5844	152	5868	166				
3 years	12	40	4	46	76.90 (63.60 to 86.39)	92.02 (80.59 to 96.97)†	98.12 (97.74 to 98.43)	98.00 (97.61 to 98.33)
5 years	30	46	20	46	60.37 (49.03 to 70.70)	69.40 (57.30 to 79.31)	98.22 (97.85 to 98.52)	97.99 (97.61 to 98.32)
8 years	46	47	45	48	49.70 (39.68 to 59.73)	50.24 (40.20 to 60.26)	98.23 (97.86 to 98.54)	98.02 (97.63 to 98.35)
10 years	57	49	50	48	44.71 (35.55 to 54.24)	47.08 (37.44 to 56.95)	98.27 (97.90 to 98.57)	98.02 (97.63 to 98.35)
HPV								
Baseline status‡:	5439	427	5210	562				
3 years	2	49	3	45	96.08 (85.62 to 99.02)†	93.74 (82.30 to 97.97)	93.50 (92.85 to 94.10)	90.97 (90.22 to 91.67)
5 years	8	67	5	58	89.34 (80.10 to 94.58)†	92.05 (82.30 to 96.65)	93.80 (93.16 to 94.39)	91.18 (90.43 to 91.87)
8 years	16	76	12	76	82.01 (72.67 to 88.66)†	86.05 (77.03 to 91.90)	93.96 (93.32 to 94.54)	91.49 (90.75 to 92.18)
10 years	24	82	13	80	75.93 (66.65 to 83.27)†	85.65 (76.84 to 91.48)	94.07 (93.43 to 94.65)	91.58 (90.84 to 92.26)
Cytology and HPV								
Baseline status§:	5365	663	5132	931				
3 years	1	51	1	50	98.07 (87.53 to 99.73)	98.03 (87.32 to 99.72)	89.76 (88.97 to 90.51)	85.35 (84.43 to 86.22)
5 years	6	70	3	64	92.07 (83.46 to 96.40)	95.51 (86.97 to 98.54)	90.05 (89.26 to 90.79)	85.55 (84.63 to 86.42)
8 years	14	79	10	84	84.25 (75.17 to 90.43)	88.98 (80.73 to 93.97)	90.19 (89.40 to 90.92)	85.85 (84.94 to 86.71)
10 years	21	86	11	88	78.89 (69.83 to 85.78)	88.42 (80.28 to 93.47)	90.31 (89.53 to 91.04)	85.92 (85.01 to 86.78)

*Women with unsatisfactory baseline smears and women with no baseline smears excluded.

†P value for test of difference of proportions, sensitivity of cytology in control arm at 3 years versus HPV testing in intervention arm at 3 (P=0.1934), 5 (P=0.4871), 8 (P=0.0255), and 10 years (P=0.0010).

‡Women with unknown HPV baseline test excluded.

§Women with both unknown HPV and unknown cytology baseline tests excluded.

Table 3| Longitudinal test characteristics using cervical intraepithelial neoplasia grade 2 or worse (CIN2+) as outcome at 3, 5, 8, and 10 years of follow-up, by cytology, human papillomavirus (HPV) status, and cytology and HPV for study arms combined and separately

Screening Test	Cumulative count of outcomes by baseline status				Negative predictive value (%)		Positive predictive value (%)	
	Intervention arm		Control arm		Intervention arm	Control arm	Intervention arm	Control arm
	Negative at baseline	Positive at baseline	Negative at baseline	Positive at baseline				
Cytology								
Baseline status*:	5844	152	5868	166				
3 years	27	64	12	73	99.54	99.79	42.28	44.07
5 years	56	70	35	75	99.03	99.39	46.34	45.60
8 years	85	71	68	77	98.48	98.77	47.05	46.63
10 years	101	73	86	77	98.13	98.38	48.73	46.57
HPV								
Baseline status†:	5439	427	5210	562				
3 years	7	83	9	72	99.87	99.83	19.51	12.84
5 years	17	108	13	91	99.68	99.75	25.63	16.30
8 years	34	121	23	113	99.33	99.54	29.02	20.72
10 years	45	128	32	121	99.07	99.32	31.12	22.57
Cytology and HPV								
Baseline status‡:	5365	663	5132	931				
3 years	3	88	5	81	99.94	99.90	13.32	8.72
5 years	11	115	9	102	99.79	99.82	17.53	11.03
8 years	26	131	19	127	99.48	99.61	20.21	14.03
10 years	37	138	28	136	99.21	99.38	21.56	15.27

*Women with unsatisfactory baseline smears and women with no baseline smears excluded.

†Women with unknown HPV baseline test excluded.

‡Women with both unknown HPV and unknown cytology baseline tests excluded.

Table 4| Longitudinal test characteristics using cervical intraepithelial neoplasia grade 3 or worse as outcome at 3, 5, 8, and 10 years of follow-up, by cytology, human papillomavirus (HPV) status, and cytology and HPV for study arms combined and separately

Screening test	Cumulative count of outcomes at follow-up by baseline status				Negative predictive value (%)		Positive predictive value (%)	
	Intervention arm		Control arm		Intervention arm	Control arm	Intervention arm	Control arm
	Negative at baseline	Positive at baseline	Negative at baseline	Positive at baseline				
Cytology								
Baseline status*:	5844	152	5868	166				
3 years	12	40	4	46	99.79	99.93	26.44	27.90
5 years	30	46	20	46	99.48	99.65	30.55	27.90
8 years	46	47	45	48	99.18	99.18	31.31	29.26
10 years	57	49	50	48	98.94	99.07	33.00	29.26
HPV								
Baseline status†:	5439	427	5210	562				
3 years	2	49	3	45	99.96	99.94	11.52	8.02
5 years	8	67	5	58	99.85	99.90	15.95	10.40
8 years	16	76	12	76	99.68	99.76	18.33	14.01
10 years	24	82	13	80	99.50	99.73	20.13	14.95
Cytology and HPV								
Baseline status‡:	5365	663	5132	931				
3 years	1	51	1	50	99.98	99.98	7.72	5.38
5 years	6	70	3	64	99.89	99.94	10.70	6.93
8 years	14	79	10	84	99.72	99.79	12.21	9.33
10 years	21	86	11	88	99.55	99.77	13.55	9.89

*Women with unsatisfactory baseline smears and women with no baseline smears excluded.

†Women with unknown HPV baseline test excluded.

‡Women with both unknown HPV and unknown cytology baseline tests excluded.

Figures

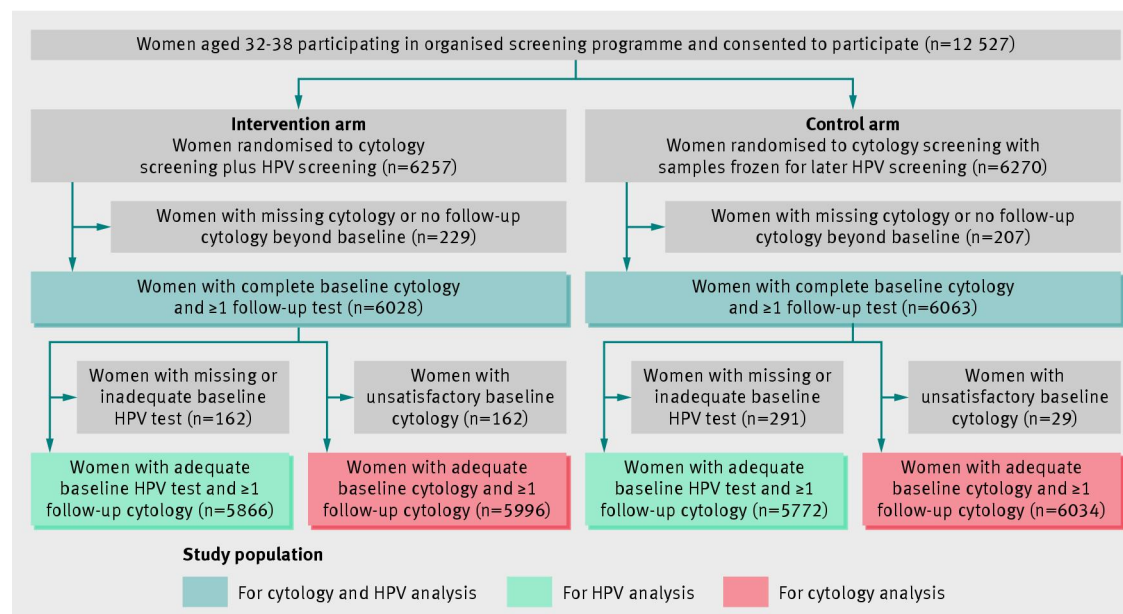


Fig 1 Study population and analyses by group. HPV=human papillomavirus

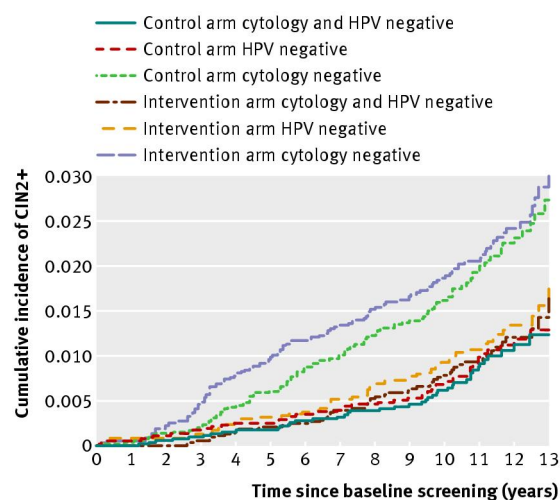


Fig 2 Cumulative incidence of cervical intraepithelial neoplasia grade 2 or worse (CIN2+) over 13 years of follow-up by study arm and baseline test result (test result groups not mutually exclusive). HPV=human papillomavirus

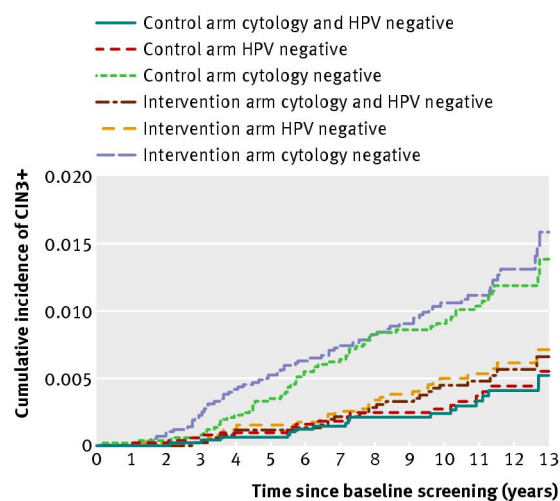


Fig 3 Cumulative incidence of cervical intraepithelial neoplasia grade 3 or worse (CIN3+) over 13 years of follow-up by study arm and baseline test result (test result groups not mutually exclusive). HPV=human papillomavirus